DOCUMENT RESUME

ED 091 341 SP 008 008

AUTHOR Heath, Robert W.; Nielson, Mark A.
TITLE The Myth of Performance-Based Teacher Education. Teacher Education Division Publication Series.
INSTITUTION Far West Lab. for Educational Research and Development, San Francisco, Calif.; NOMOS Inst., Berkeley, Calif.
REPORT NO R-A73-10
PUB DATE Feb 73
NOTE 23p.; Paper presented at the Annual Meeting of the American Educational Research Association (New Orleans, Louisiana, February 1973)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS *Academic Achievement; *Definitions; Educational Research; *Performance Based Teacher Education; *Performance Criteria; Problems; *Research; *Research Design; Teacher Behavior; Teacher Characteristics

ABSTRACT

This study analyzed 42 of the 50 studies on performance-based teacher education cited by Rosenshine and Furst and the teaching variables proposed. The studies were analyzed for seven characteristics: a) criterion of student achievement, b) operational definitions of teaching behavior, c) statistical results, d) legitimacy of statistical tests, e) sample characteristics, f) study conditions, and g) comparability of groups. The results of this analysis led to the conclusion that research literature on the relation between teacher performance and student achievement does not offer an empirical basis for the prescription of teacher-training objectives. This is due to sterile operational definitions of both teaching and achievement and fundamentally weak research designs in the cited studies. Moreover, given the well-documented strong association between student achievement and variables such as socioeconomic status and race, the effects of techniques of teaching on achievement (as these variables are conventionally defined) are likely to be inherently trivial. (JA)

ED 091341

NOMOS INSTITUTE
BOX 941
BERKELEY, CALIFORNIA 94701
415 / 451-3072

THE MYTH OF PERFORMANCE-BASED TEACHER EDUCATION

Robert W. Heath

and

Mark A. Nielson

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

A Paper Presented to
the American Educational Research Association
February 1973, New Orleans

SP008 008

THE MYTH OF PERFORMANCE-BASED TEACHER EDUCATION

In recent years research and development in teaching has been strongly influenced by the idea of "performance-based teacher education." That is, the identification of teacher skills that are related to increased student achievement and the corresponding development of teacher training programs and materials (see Gage, 1972). Since hundreds of millions of dollars annually and substantial human resources are involved, the scientific basis for this movement is of considerable importance. In our opinion, an analysis of the research on the relation between teacher performance and student achievement fails to reveal an empirical basis for performance-based teacher education.

The contribution by Rosenshine and Furst, "Research on Teacher Performance Criteria" in the book Research in Teacher Education: A Symposium, 1971 (edited by B.O. Smith), reviews a carefully selected set of studies. Rosenshine's several writings on this and related topics establish him as knowledgeable of this body of literature.

Rosenshine and Furst cite some 50 studies and propose 11 teacher behavior variables which they regard as "the most promising of the variables studied." The variables proposed are (1) clarity, (2) variability, (3) enthusiasm, (4) task-oriented and/or business-like behaviors, (5) student opportunity to learn criterion material, (6) use of student ideas and general indirectness, (7) criticism, (8) use of structuring comments, (9) types of questions, (10) probing, and (11) level of difficulty of instruction, in that order.

The studies cited, and the teaching variables proposed, were carefully selected to represent the most conclusive and the "best" research

to be found. Rosenshine and Furst say:

> The variables described below were abstracted from more detailed reviews (Rosenshine, 1970 a,b) and are the most promising of the variables studied. The variables with the strongest support are presented first.
>
> A description of the eleven strongest variables contained in this research is presented below together with the results obtained when these variables were studied. The best results were obtained on the first five variables; the results were less conclusive on the last six variables. The major focus is upon process-product studies, but experimental classroom studies are cited whenever appropriate.
>
> ...Unlike the myriad competencies outlined by the model program builders, the skills listed below have some support in studies of naturally occurring teacher behavior and student achievement.

The studies included in the Rosenshine and Furst review range from major long-term research efforts to minor studies of very short duration.

We have carefully analyzed 42 of the studies cited for seven characteristics: (1) criterion of student achievement; (2) operational definitions of teaching behavior; (3) statistical results; (4) legitimacy of statistical tests; (5) sample characteristics; (6) study conditions; and (7) comparabiltiy of groups.

Method

The method for gathering the information contained in the operational-definition table (Table 1) and our criteria for determining significance claims follow.

In searching for the author's operational definition of the teacher behavior under scrutiny, we referred to the manner in which Rosenshine and Furst cite the study. Often, Rosenshine and Furst quote the author's operational definition of the variable. For example, under clarity: "The investigators used different descriptions

TABLE 1. OPERATIONAL DEFINITIONS OF TEACHER-PERFORMANCE AND FACE VALIDITY DECISIONS

VALIDITY	
	CLARITY
YES	"Clarity of Presentation" on the Stanford Teacher Competence Appraisal Guide (seven-point rating scale). (Belgard, Rosenshine, & Gage, 1968)
NO	"Difficulty of Lesson". "Evaluates the lesson in terms of its ease or difficulty for the children." (Chall & Feldmann, 1966)
NO	"Disorganized-Systematic" on an adjective checklist. "Organization of Lesson" and "Lesson Presentation" on the Stanford Teacher Competence Appraisal Guide (seven-point rating scale). (Fortune, 1967)
YES	"Clarity of Presentation" on the Stanford Teacher Competence Appraisal Guide (seven-point rating scale). (Fortune, Gage, & Shutes, 1965)
YES	"Were the points he (she) made clear and easy to understand?" "Were his (her) presentations usually well-organized?" "Coherence": the teacher's "articulateness, his ability to make points which the students understand." (Solomon, Bezdek, & Rosenberg, 1963)
YES	Q-sort on the "Intellectually Effective vs. Intellectually Ineffective" continuum. "The intellectually effective teacher . . . was the one who appeared to be able to explain concepts clearly and such that the students seemed to be gaining understanding." (Wallen, 1966)
	CLARITY: LOW-INFERENCE VARIABLES
YES	"Vagueness Categories," including "ambiguous designation," "approximation," "indeterminate quantification," and "probability." (Hiller, Fisher, & Kaess, 1969)
NO	"Interpretation," a category for analysis of taped lessons. (Solomon, Bezdek, & Rosenberg, 1963)
NO	"Percent of questions leading to pupil response." (Wright & Nuthall, 1970)
	CLARITY: ORGANIZATION
YES	"A typical item on the Disorganization scale is, 'There are long periods during which the class does nothing.'" The study focused on "the socio-emotional climate of the class" and on "using student perceptions of it as predictors of class learning." (Anderson & Walberg, 1968)
YES	"Organization of the Lesson" on the Stanford Teacher Competence Appraisal Guide (seven-point rating scale). (Belgard, Rosenshine, & Gage, 1968)
YES	"Disorganized-Systematic" on an adjective checklist. "Organization of Lesson" and "Lesson Presentation" on the Stanford Teacher Competence Appraisal Guide (seven-point rating scale). (Fortune, 1967)
YES	"Organization of the Lesson", an item on the Stanford Teacher Appraisal Guide, was not studied. (Fortune, Gage, and Shutes, 1966)
YES	"Organization", if the teacher tries "to relate ideas, theories, current events, etc. in an ordered, meaningful, coherent, and logical fashion." (Solomon, Bezdek, & Rosenberg, 1963)
YES	"Disorganization", a four-point scale on the Learning Environment Inventory, "a measure of the social environment of learning." (Walberg, 1969)
YES	"Disorganized" on the Classroom Climate Questionnaire. (Walberg & Anderson, 1968)
	VARIABILITY
YES	Listed teacher aids (e.g., charts, film, radio, etc.) and pupil materials (e.g., magazines, paint, workbook, etc.) (Anthony, 1967)
NO	"Inflexible-Adaptable" on an adjective checklist. (Fortune, 1967)
NO	"Multiple Cognitive Levels": "Ratio of number of lines devoted to analytic and evaluative substantive-logical processes to number of lines devoted to empirical logical process." (Furst, 1967)
YES	Answers to "Did the instructor follow up new and unexpected topics, suggested by student questions or suggestions?" and "Did he (she) sometimes make changes in the way material was being presented?" (Solomon, Bezdek, and Rosenberg, 1963)
NO	"Verbosity": "Frequency of verbal output." (Thompson and Bowers, 1968)
NO	"This school has a great deal of resource materials for extra study" and "This school offers extra learning facilities which include a library, audiovisual aids, etc." (Torrance and Parent, 1966)
NO	A Biographical Inventory including "books in home," "non-school reading," "musician," "career discussion with adults," etc. (Walberg, 1969)
	VARIABILITY: FLEXIBILITY
YES	Comparison of ratio of Indirect Influence (I) to Direct Influence (D) across activity periods. (Flanders, 1970)
YES	"Flexibility": ". . . beginning with largest cell frequencies, a count of cells necessary to account for 60 percent of the tallies." (Soar, 1966)
YES	Flexibility, "The number of times a teacher changes behavior in a given time interval." (Vorreyer, 1965)
	ENTHUSIASM
NO	"Dull-Stimulating" on an adjective checklist. (Fortune, 1966)
NO	"Dull-Stimulating" and "Apathetic-Alert" on an adjective checklist. (Fortune, 1967)
NO	"Apathetic-Alert," "Dull-Stimulating," and "Stereotyped-Original" on an adjective checklist. (Kleinman, 1964)
YES	Teachers were asked to lecture in either an enthusiastic or unenthusiastic manner. (Mastin, 1963)
YES	"Energy" and "Enthusiasm" on a teacher behavior rating scale. "Was the instructor enthusiastic and excited about the subject matter?" on a student questionnaire. (Solomon, Bezdek, & Rosenberg, 1963)
NO	"Dull vs. Stimulating" on a teacher behavior rating scale. (Wallen, 1966)

TABLE 1. (Cont.)

VALIDITY	TASK-ORIENTED AND/OR BUSINESSLIKE BEHAVIORS
NO	"Task-oriented, teacher-centered" (e.g., "The teacher tells us exactly how to do our homework") vs "Task-oriented, pupil-centered" (e.g., "The teacher has us help each other in class") on a pupil survey. (Beiderman, 1964)
NO	Rating scales for "approach to learning, information" and "approach to learning; thinking" on an observation inventory. (Chall & Feldmann, 1966)
NO	"Evading-Responsible," "Erratic-Steady," "Disorganized-Systematic," and "Excitable-Poised" on an adjective checklist. (Fortune, 1967)
NO	"Evading-Responsible," "Erratic-Steady," Excitable-Poised," and "Disorganized-Systematic" on an adjective checklist. (Kleinman, 1964)
YES	Student ratings on "My teacher has encouraged this class to think of original solutions to mathematical problems," "In this class we have one of the most hard-working classes in the school," and "My teacher has encouraged this class to work out and have our own answers to problems." (Torrance & Parent, 1966)
YES	"Achievement Oriented vs. Non-Achievement Oriented" on an observer Q-sort. (Wallen, 1966)
YES	Student teachers' grades in student teaching depended on their students' achievement scores. (Wittrock, 1962)
	STUDENT OPPORTUNITY TO LEARN CRITERION MATERIAL
YES	"Substantive," defined as "subject matter of instruction." (Bellack, Kliebard, Hyman, & Smith, 1966)
YES	"Total number of relevant words" in lecture. (Rosenshine, 1968)
YES	Whether or not the lesson answered particular test items ("content relevance"). (Shutes, 1969)
	USE OF STUDENT IDEAS AND GENERAL INDIRECTNESS: USE OF STUDENT IDEAS
YES	"Accepts or uses ideas of pupils" on an observation record. (Flanders, 1970)
YES	"Accepts or uses student's answer or idea" on observation record. (Perkins, 1965)
YES	"Accepts or uses ideas of student" on observation record. (Soar, 1966)
NO	"Repetition of response" in tape analysis. (Wright & Nuthall, 1970)
	USE OF STUDENT IDEAS AND GENERAL INDIRECTNESS: INDIRECTNESS
YES	"Accepts or uses ideas of pupils" on an observation record. (Flanders, 1970)
NO	"High" or "Low" in emotional climate on an observation record (High = warmth and friendliness). (Medley & Mitzel, 1959)
	USE OF STUDENT IDEAS AND GENERAL INDIRECTNESS: I/D RATIO
YES	Ratio of "Indirect Influence" to "Direct Influence." (LaShier, 1967)
	USE OF STUDENT IDEAS AND GENERAL INDIRECTNESS: EXPERIMENTAL STUDIES
YES	Teachers were trained to create a "positive social-emotional climate" using Flanders I. A. (Carline, 1969)
YES	The experimental group was trained to be more indirect in their teaching. (Gunnison, 1968)
YES	"Teacher-Centered" vs. "Pupil-Centered" treatments. (Herman, Potterfield, Dayton, & Amershek, 1969)
YES	Teachers were asked to be either "Directive" or "Responsive." (Miller, 1966)
	CRITICISM
YES	"Intensity of negative support observed." (Anthony, 1967)
YES	". . . bawling someone out . . ." in classroom observation record. (Cook, 1967)
YES	"Criticizing or justifying authority" on observation record. (Flanders, 1970)
YES	"Negative motivational statements" on observation record. (Harris, Morrison, Serwer, & Gold, 1968)
NO	"Teacher Control," "A low degree of pupil disruptive acts during observation." (Hunter, 1969)
NO	"Threatens, warns" on observation checklist. (Morsh, 1956)
YES	". . . rejection or correction of student's response . . ." and "criticizes or justifies authority . . ." on observation record. (Perkins, 1965)
YES	"Steady-state criticism" and "Pupil initiation following teacher criticism" on observation record. (Soar, 1966)
YES	"Manifest teacher disapproval" and "Emotional tone or mood of teacher" on observation record. (Spaulding, 1965)
YES	"Disparaging vs. non-disparaging" on rating scale and Q-sort. (Wallen, 1966)
NO	"Affirmative and Negative" comments on observation record. (Wright & Nuthall, 1970)
	USE OF STRUCTURING COMMENTS
YES	"Beginning the Lesson" and "Ending the Lesson" on the Stanford Teacher Competence Appraisal Guide (a seven-point rating scale). (Belgard, Rosenshine, & Gage, 1968)
YES	"Beginning the Lesson" and "Ending the Lesson" on the Stanford Teacher Competence Appraisal Guide (a seven-point rating scale). (Fortune, 1967)
YES	"Teacher structure," "Deviation of individual teacher's amount of lines devoted to structuring." (Furst, 1967)
YES	Number of "Verbal markers of importance" in lesson transcript. (Pinney, 1969)
NO	No variables resembling "structuring" could be located. (Soar, 1966)
YES	"Structuring (total)," "teacher information following question," and "terminal structuring." (Wright & Nuthall, 1970)

TABLE 1. (Cont.)

VALIDITY	TYPES OF QUESTIONS
NO	"Communication," "Management," and "Encouragement" were observed in classroom. (Conners & Eisenberg, 1966)
YES	Questions requiring pupil understanding of the meaning vs. those requiring only recognition of the symbol involved. (Harris, Morrison, Serwer, & Gold, 1968)
YES	Seven question categories: neutral, rhetorical, factual, clarifying, associative, critical thinking, and values. (Kleinman, 1964)
YES	"Asks questions about content . . ." vs. "Asks questions that stimulate thinking . . ." (Perkins, 1964)
YES	Questions were categorized as "organizing," "hypothetical," "opinion," "factual," "interpretation," or "personal reference." (Solomon, Bezdek, & Rosenberg, 1963)
YES	"Technique – eliciting a verbal response in an open-ended fashion," and "Technique – eliciting a specific or solution that the teacher has in mind" on a rating scale. (Spaulding, 1965)
NO	"Convergence-divergence" on an observation record. (Thompson & Bowers, 1968)
YES	"Closed questions" and "open questions" on an observation record. (Wright & Nuthall, 1970)
	PROBING
YES	"Teacher encourages interpretation, generalization, solution" on an observation record. (Soar, 1966)
YES	"Technique – eliciting clarification in a non-threatening way" on an observation record. (Spaulding, 1965)
YES	"Redirects question," "reciprocates to lift," and "reciprocates to extend" on an observation record. (Wright & Nuthall, 1970)
	LEVEL OF DIFFICULTY OF INSTRUCTION
YES	Students rated teachers in hypothetical situations as using "most strict standard," "most liberal standard," or "somewhere in between." (Nikoloff, 1965)
NO	The variable referred to by Rosenshine & Furst could not be located. (Torrance & Parent, 1966)
YES	"Difficulty" on Learning Environment Inventory. (Walberg, 1969)

of clarity:

1. 'Clarity of Presentation' (Belgard, Rosenshine, and Gage, 1968; Fortune, 1967; Fortune, Gage, and Shutes, 1966)...." (p. 44). If this was the case, we simply quote this definition from the text of the study and note it in the table.

In many other cases, Rosenshine and Furst merely describe the author's operational definition. For example, on enthusiasm: "Teacher enthusiasm has been assessed by...observer estimation of the amount of vigor and power exhibited by the teacher during classroom presentation...." (p. 46). In these cases, we attempted to locate the teaching variables in each study which most closely corresponded to Rosenshine and Furst's description.

In a few cases, we were unable to locate the teaching variable described by Rosenshine and Furst. These cases have been noted in the table.

In deciding whether or not the author claimed significant results, we concentrated on the tables and discussion of results in each study. If the results table(s) specifically indicated significance levels for the variable in question, we regarded this as a claim of significance. If indications of significance were absent from the tables, but were present in the discussion of results, we regarded this as a claim of significance as well. In several studies, an explicit claim of significance was absent from both tables and discussion of results, but the discussion conveyed the impression that the author was claiming significance. Thus, we had three criteria for concluding that the author claimed significance: (1) whether or not he reported significance levels in tables of results, (2) whether or not he claimed significance in the text of his report, or (3) whether or not the significance

of results was implicit in his discussion. If a study met any one of these cri[illegible] we concluded that the author was claiming significance.

Simil[illegible]teria were used in the cases where we concluded that the author did not claim significant results. We reached this conclusion when the relevant values in the results tables did not reach significance levels specified in the tables, when the author reported non-significant results in his discussion, or when the non-significance of results was implicit in his discussion.

Clearly, there is a difference between the statement that the author claimed significant results and the statement that the author obtained significant results. The fact that we were concerned with the former statement means that we were more stringent in our judgments concerning the significance of results than were Rosenshine and Furst. In the cases where Rosenshine and Furst report that the author found significant results, but we concluded that the author did not claim significance, Rosenshine and Furst looked up the significance of the reported values themselves. In view of other design considerations, we believe that this is an unwarranted procedure. Most of the studies, for instance, failed to report the tenability of critical assumptions. If these assumptions were not met, the reported values in these studies could be spuriously high or low. The original authors are likely to have very good reasons for not claiming significance. These reasons are frequently not made explicit in the reports.

The criteria used for analyzing the design and methodology of the studies are explained in detail in the Technical Appendix of our larger report, "An Analysis of Selected Studies of Teacher Effectiveness." Each study was analyzed for such research design features as random assignment, number of teachers and number of students per teacher,

adjustment of post-test scores for initial differences, experimental independence of variables, and significance or non-significance claims. The studies were also analyzed for the legitimacy of their statistical procedures: assumptions of linear relationships, parallel regression slopes, normality, homogeneity of variance, method of estimating the slope used in adjusting post-test scores, and the effect of extreme data points ("outliers").

Results

The results of the analysis are summarized in Table 1 and Table 2. Table 1 lists each author's operational definition of the teaching variables cited by Rosenshine and Furst. Our judgment concerning the appropriateness of the operational definition to the variable cited is indicated. In our judgment, 26 of the 84 operational definitions of teacher behavior do not correspond to the variable cited.

Table 2 lists the studies, the variables for which they are cited in Rosenshine and Furst's review, and a number of their basic characteristics.

Fifty-eight citations used the teacher as the unit of sampling (79%) and 10 (13%) used the pupil. In the other 10 cases (13%), the information reported was not sufficient to determine the unit of sampling.

One citation reported a linear relation (1%) and 61 did not report having tested the assumption of linearity (78%). Linearity is not relevant to the remaining 16 citations (21%).

Five citations report randomly assigned subjects to treatments (6%) and 65 did not randomize (83%). Eight did not report enough information for a judgment (11%).

TABLE 2. DESIGN AND METHODOLOGY CHARACTERISTICS

STUDY	TEACHER BEHAVIOR VARIABLE	TEACHER AS SAMPLING UNIT	LINEARITY	RANDOM ASSIGNMENT	POOLED WITHIN SCORE	PARALLEL SLOPE	OUTLIERS	AT LEAST 10 CLASSES	AT LEAST 15 PER CLASS	NORMALITY	COVARIATE INDEPENDENT OF TEACHING	HOMOGENEITY OF VARIANCE	SIGNIFICANCE CLAIM
ANDERSON & WALBERG, 1963	CLARITY (ORGANIZATION)	YES		NO	NA	NA	NA	YES	NO	–	NA	–	YES
ANTHONY, 1967	VARIABILITY	YES		NO	[1]	[1]	[1]	YES	YES	–	–	–	YES
	CRITICISM	YES		NO				YES	YES	–	–	–	NO
BEIDERMAN, 1964	TASK ORIENTED BEHAVIORS	YES	–	NO	NA[4]	NA[4]	NA[4]	YES	YES	–	–	–	NO
BELGARD, ROSENSHINE, & GAGE, 1968	CLARITY	YES		NO	NO			YES	YES	–	–	–	NO
	STRUCTURING COMMENTS	YES		NO	NO			YES	YES	–	–	–	NO
BELLACK, KLIEBARD, HYMAN, & SMITH, 1966	OPPORTUNITY TO LEARN	NA	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NO[5]
CARLINE, 1969	USE OF STUDENT IDEAS	YES	NA[5]	NO	NA[5]	NA[5]	NA[5]	YES	YES	–	NA[5]	–	NO
CHALL & FELDMANN, 1966	CLARITY	NO		NO	NO			NA	NA	–	–	–	NO
	TASK ORIENTED BEHAVIORS	NO	–	NO	NO			NA	NA	–	–	–	NO
CONNERS & EISENBERG, 1966	TYPES OF QUESTIONS	YES		NO	NO			YES	YES	–	–	–	YES
COOK, 1967	CRITICISM	NA	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NO
FLANDERS, 1970	VARIABILITY (FLEXIBILITY)	YES	–	NO	NO	–	–	YES	YES	–	NO	–	NO
	USE OF STUDENT IDEAS	YES		NO	NO			YES	YES	–	NO	–	NO
	CRITICISM	YES	–	NO	NO			YES	YES	–	NO	–	NO
FORTUNE, 1966	ENTHUSIASM	YES	–	NO	NO	–	–	YES	NO	–	NO	–	NO
FORTUNE, 1967	CLARITY	YES		NO	NO	–	–	YES	YES	–	YES	–	NO
	VARIABILITY	YES		NO	NO	–	–	YES	YES	–	YES	–	NO
	ENTHUSIASM	YES	–	NO	NO	–	–	YES	YES	–	YES	–	NO
	TASK ORIENTED BEHAVIORS	YES	–	NO	NO	–	–	YES	YES	–	YES	–	NO
	STRUCTURING COMMENTS	YES	–	NO	NO	–	–	YES	YES	–	YES	–	NO
FORTUNE, GAGE, & SHUTES, 1966	CLARITY	YES		YES	NO	–	–	YES	NO	–	NO	–	NO
FURST, 1967	VARIABILITY	NA		NO	[2]	[2]	[2]	YES	YES	[2]	[2]	[2]	YES
	STRUCTURING COMMENTS	NA	–	NO	[2]	[2]	[2]	YES	YES	[2]	[2]	[2]	YES
GUNNISON, 1968	USE OF STUDENT IDEAS	NO	YES	NO			–	NO	YES	–	YES	–	YES
HARRIS, MORRISON, SERWER, & GOLD, 1968	CRITICISM	YES		NO	NO	–	–	YES	YES	–	–	–	YES
	TYPES OF QUESTIONS	YES		NO	NO	–	–	YES	YES	–	–	–	NO
HERMAN, POTTERFIELD, DAYTON, & AMERSHEK, 1969	USE OF STUDENT IDEAS	YES	–	YES			–	NO	YES	–	NO	–	NO
HILLER, FISHER, & KAESS, 1969	CLARITY	YES		NO	NO			YES	YES	–		–	YES
HUNTER, 1969	CRITICISM	YES		NO	NO	–	–	NO	NO	–	NO		YES
KLEINMAN, 1964	ENTHUSIASM	NA	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NO[5]
	TASK ORIENTED BEHAVIORS	NA	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NO
	TYPES OF QUESTIONS	NO	NA	NO	NA	NA	NA	YES	YES	–	NA	–	YES
LaSHIER, 1967	USE OF STUDENT IDEAS	YES	NA[10]	NO	NA[10]	NA[10]	NA[10]	NO	YES		NA[10]	NA[10]	YES
MASTIN, 1963	ENTHUSIASM	NO	NA[5]	NO	NA[5]	NA[5]	NA[5]	YES	YES	–	NA[5]	–	YES
MEDLEY & MITZEL, 1959	USE OF STUDENT IDEAS	YES	–	NO	YES			YES	YES	–		–	NO
MILLER, 1966	USE OF STUDENT IDEAS	YES		YES			–	NO[12]	NO	–	–	–	NO
MORSH, 1956	CRITICISM	YES			NO			YES					NO
NIKOLOFF, 1965	LEVEL OF DIFFICULTY	NO	NA	NO	NA	NA	NA	YES	YES	–	NA		NO
PERKINS, 1964	TYPES OF QUESTIONS	NO	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NO
PERKINS, 1965	USE OF STUDENT IDEAS	NA	NA[5]	NO	NA[5]	NA[5]	NA[5]	NO					NO[5]
	CRITICISM	NA	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NO[5]
PINNEY, 1969	STRUCTURING COMMENTS	YES		YES	YES	–	–	YES	YES	–	YES	–	YES
ROSENSHINE, 1968	OPPORTUNITY TO LEARN	YES		NO	NO		–	YES	YES	–		–	YES
SHUTES, 1969	OPPORTUNITY TO LEARN	YES		YES	YES	YES	–	YES	YES	–	YES	–	YES
SOAR, 1966	VARIABILITY (FLEXIBILITY)	YES		NO	YES	–	–	YES	YES	–	NO	–	NO
	USE OF STUDENT IDEAS	YES		NO	YES		–	YES	YES	–		–	NO
	CRITICISM	YES	–	NO	YES	–	–	YES	YES	–	YES	–	YES
	STRUCTURING COMMENTS	YES	–	NO	YES	–	–	YES	YES	–	YES	–	NO
	PROBING	YES		NO	YES	–	–	YES	YES	–	YES	–	YES
SOLOMON, BEZDEK, & ROSENBERG, 1963	CLARITY	YES		NO	NO	–	–	YES	YES	–	NO	–	YES
	VARIABILITY	YES	–	NO	NO	–	–	YES	YES	–	NO	–	NO
	ENTHUSIASM	YES	–	NO	NO	–	–	YES	YES	–	NO	–	YES
	TYPES OF QUESTIONS	YES		NO	NO		–	YES	YES	–	NO	–	NO
SPAULDING, 1965	CRITICISM	YES		NO	NO			YES	YES	–		–	YES
	TYPES OF QUESTIONS	YES		NO	NO	–	–	YES	YES	–		–	YES
	PROBING	YES		NO	NO			YES	YES	–	YES	–	YES
THOMPSON & BOWERS, 1968	VARIABILITY	NO	NA[5]	NO	NA[5]	NA[5]	NA[5]	YES	YES	–	NA[5]	–	NO
	TYPES OF QUESTIONS	NO	NA[5]	NO	NA[5]	NA[5]	NA[5]	YES	YES	–	NA[5]	–	YES
TORRANCE & PARENT, 1966	VARIABILITY	YES	–	NO	NO			YES	YES	–	YES	–	YES
	TASK ORIENTED BEHAVIORS	YES		NO	NO			YES	YES	–	YES	–	YES
	LEVEL OF DIFFICULTY	YES		NO	NO			YES	YES		YES		NO
VORSEYER, 1965	VARIABILITY (FLEXIBILITY)	YES	NA[5]	NO	NA[5]	NA[5]	NA[5]	NO	YES	NA[5]	NA[5]	NA[5]	NO
WALBERG, 1969	CLARITY	YES		NO	NO	–	–	YES	YES		NO	–	NO
	VARIABILITY	YES		NO	NO	–	–	YES	YES	–	NO	–	YES
	LEVEL OF DIFFICULTY	YES		NO	NO	–	–	YES	YES	–	NO	–	YES
WALBERG & ANDERSON, 1968	CLARITY	NO	–	NO	NA	NA		NA	NA	–	NO	NA	NO
WALLEN, 1966	CLARITY	YES	–	NO	YES		–	YES	YES	–		–	YES
	ENTHUSIASM	YES	–	NO	YES		–	YES	YES	–		–	YES
	TASK ORIENTED BEHAVIORS	YES	–	NO	YES			YES	YES				YES
	CRITICISM	NA	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NA[5]	NO[5]
WITTROCK, 1962	TASK ORIENTED BEHAVIORS	YES		NO	NO			NO	YES	–		–	YES
WRIGHT & NUTHALL, 1970	CLARITY	YES		NO	NO	–	–	YES	YES	–	NO	–	NO
	USE OF STUDENT IDEAS	YES	–	NO	NO			YES	YES	–	NO	–	NO
	CRITICISM	YES		NO	NO		–	YES	YES	–	NO	–	NO
	STRUCTURING COMMENTS	YES		NO	NO		–	YES	YES	–	NO	–	YES
	TYPES OF QUESTIONS	YES		NO	NO	–	–	YES	YES	–	NO	–	NO
	PROBING	YES	–	NO	NO		–	YES	YES	–	NO	–	YES

NA = NOT APPLICABLE. THE DASHES MEAN "NOT REPORTED"

1. UNADJUSTED POSTTEST MEDIANS OF GRADE EQUIVALENTS USED AS CRITERIA
2. NO INFORMATION CONCERNING THE NATURE OF THE STATISTICAL TEST IS GIVEN
3. DESIGN WAS ANOVA OF POSTTEST MEANS
4. ANALYSIS WAS RANK ORDER CORRELATION OF POSTTEST MEANS WITH TEACHING VARIABLE
5. NO STATISTICAL ANALYSIS PERFORMED FOR THIS VARIABLE
6. ANALYSIS WAS t TEST OF UNADJUSTED POSTTEST PUPIL MEANS
7. POST MEAN MINUS PRE MEAN USED AS CRITERION
8. NO STATISTICAL TESTS REPORTED
9. NO SIGNIFICANCE TEST OR CORRELATION REPORTED
10. RANK ORDER CORRELATION OF POST PRE MEDIANS
11. ANOVA OF UNADJUSTED POSTTEST MEANS
12. N = 4

Eleven studies used a pooled within-group estimate of regression slope in adjusting for initial ability (14%), while 41 did not (53%). In five cases, there was insufficient information reported to make a judgment (6%). The remaining 21 citations either did not adjust for initial ability or did not report statistical tests for the variable in question (27%).

One citation reported that there was no treatment-slope interaction (1%), 56 did not report having tested the assumption of parallel slopes for all treatment groups (72%), and 21 either did not adjust for initial ability or did not report statistical tests for the variable in question (27%).

Fifty-eight citations (74%) did not report the disposition of outliers (extreme data points). This characteristic does not apply to the remaining 20 citations (26%).

Sixty citations included at least 15 classes (77%), and eight reported less than 15 classes (10%). Ten citations did not report sufficient information for a judgment to be made (13%).

Sixty-one citations included at least 15 pupils per class (78%), and five included less than 15 per class (6%). In two cases, there was insufficient information (3%), and in the remaining 10 cases, no statistical tests were performed for the variable in question (13%).

Seventy citations did not report whether the assumption of a normal distribution had been tested (90%). The remaining eight did not report performing statistical tests for the variables in question (10%).

The covariate was found to be independent of teaching behavior in 15 studies (19%). It was not independent of teaching behavior in 22 cases (28%). There was insufficient information for judgment in

25 citations (32%). The category does not apply to the remaining 16 cases (21%).

Testing for homogeneity of variance was not reported in 68 studies (87%). The assumption does not apply to the remaining 10 cases (13%).

Thirty-three of the reports claimed a significant relation between the variable in question and some criterion measure (42%). The other 45 citations did not claim significant results (58%).

Discussion

These results lead us to conclude that an empirical basis for performance-based teacher education does not exist. The conception, design, and methodology typical of these studies precludes their use as such a basis.

The operational definitions of teaching shown in Table 1 do not correspond, in many instances, to the names of the teaching variables cited. For example, a rating of "difficulty of the lesson" (in Chall and Feldmann, 1966) is cited as the basis for a study of "clarity." Nearly a third of the operational definitions studied seem similarly inappropriate.

Further, operational definitions of teacher behavior that have little in common are often combined as examples of a single teaching variable. Note that a rating of "organization of the lesson" (Belgard, Rosenshine, and Gage, 1968) is combined with "disorganized" as a characteristic of the classroom climate (Walberg and Anderson, 1968) under the single variable "organization."

Beyond these logical difficulties, the operational definitions reported in the studies generally reflect a picture of the teacher in a traditional lecture-discussion role and are usually defined so vaguely as to be of little use in training teachers. "Clarity of

presentation" on a seven-point rating scale (Fortune, Gage, and Shutes, 1966) assumes that teaching requires a "presentation." A rating on "clarity," with no explication of what behavior "clarity" involves, does not seem useful. Continued research on such sterile definitions of teaching seems unlikely to provide a basis for training teachers.

The operational definitions of student achievement are similarly shallow. For example, the criterion of student achievement used in several studies is a ten-question multiple-choice test of information based on Atlantic magazine articles on economic, political, and social conditions in Yugoslavia published between November 1964 and August 1965.

The research design of most of these studies is fundamentally weak. Only five of the 78 citations report using random assignment of pupils to treatments, yet they employ statistical analyses whose interpretation implies randomization. As Lord (1963) has said:

> If the individuals are not assigned to the treatments at random, then it is not helpful to demonstrate statistically that the groups after treatment show more difference than would be expected by random assignment....

In most (45 of 78) of the studies cited, the author did not claim to find a significant relation between the teacher-performance variable and student achievement. Further, in many of these studies negative or clearly non-significant relations are reported. These negative results are ignored.

Though most studies employed adjusted post-test means as criteria, only about 10% used the preferred method of adjustment.

Aside from questions of research design, the statistical analyses reported make interpretations of the studies extremely undependable. Critically important assumptions such as linearity, normality, homogeneity of variance, and parallel slopes are almost universally

unreported (see Table 2). There is reason to believe that in many instances these assumptions are not, in fact, tenable. With no information, one might choose to believe that a particular basic assumption was met, or could somehow be disregarded in a particular study. The compound effect of unreported and unmet assumptions is another matter. For example, homogeneity of variance and normality are not independent assumptions; if one is not met, the other becomes even more critical. One cannot, with scientific integrity, dismiss such assumptions and interpret the results at will.

One of the pitfalls commonly encountered in this literature is the item-factor fallacy in factor analysis. Rosenshine and Furst report significant results in a small number of studies employing factor analysis. When a relevant item of teacher-behavior loaded on a factor that was significantly correlated to achievement, they claim a significant relationship between the item and achievement. A significant correlation between a factor and student achievement, however, does not necessarily imply significant correlations between achievement and every teacher-behavior item loading on that factor.

The performance-based-teacher-education model does not treat two important types of variables. It ignores what is to be taught. Though the studies reviewed here were concerned with everything from aircraft mechanics to reading, no effort is apparent in identifying the possible interactions between teacher-behavior variables and content. Is there one set of best teaching behaviors for teaching everything? If there is an important interaction between content and teaching behavior (given cognitive achievement as criteria), then the conclusions about what behavior is effective may be determined as much by content as by teacher behavior.

The model ignores who is to be taught. Despite persistent evidence that variables such as socio-economic status and race are more important determinants of achievement level than teacher behavior, the research on teacher-behavior variables largely ignores such differences among students. Similarly, the studies cited by Rosenshine and Furst cover a wide student age range (pre-school to adult), yet the idea that effective teacher behavior might be different for different age groups is ignored when conclusions are drawn from such collections. Is there one set of best teacher behaviors for teaching everything to everybody?

Other Reviews

The studies cited by Rosenshine and Furst were selected as the best and most conclusive available. Since our analysis and interpretation of these studies leads us to doubt the case for demonstrated relations between teacher-behavior variables and student achievement, it seems prudent to look at the conclusions that have been stated by others who have reviewed the research on this subject.

In 1958, Brim (1958, p. 32) concluded, after reviewing the status of the research:

> However, although there is a vast body of research on the relation of teacher characteristics to effectiveness in teaching, the reviews of this research...show no consistent relation between any characteristics, including intelligence, and such teaching effectiveness.

In 1967, Stephens (1967, p. 83-84) concluded:

> It is always possible that the vast number of investigations have "just happened" to hit on the more intransient areas of growth. It is also possible that they have just happened to miss the areas which would have shown the effects of the device being tested. But these possibilities are not very comforting. The argument implies that some systematic force has been at work to keep our sample a

biased sample. The argument rests on the assumption that the unmeasured areas of character growth are more responsive to the devices being tested than the more measurable areas of academic achievement. Actually, that assumption is very dubious, a point which will be developed at some length in Chapter 9.

The other four arguments are all confronted by one very serious obstacle. The fact is, insensitive as the tests may be and over-controlled or under-controlled as some experiments probably are and exacting as standards undoubtedly are, a great deal of growth does appear and does meet the standards. The investigations cited do not fail to reveal growth. They merely fail to reveal differences in growth attributable to the administrative (teaching) variables. If we use other variables, such as background factors, moreover, marked differences in growth also come through. If the tests, and the designs, and the criteria of significance permit such differences to appear, it is difficult to see why they should not also permit differences in administrative (teaching) factors to come through if these were present.

In 1968, Dubin and Taveggia (1968, p. 35) concluded, after reviewing 91 comparative studies of college teaching technologies conducted between 1924 and 1965:

> These data demonstrate clearly and unequivocally that there is no measurable difference among truly distinctive methods of college instruction when evaluated by student performance on final examinations.

Mood (1970, p. 7) concluded:

> We can only make the not very useful observation that <u>at the present moment we cannot make any sort of meaningful quantitative estimate of the effect of teachers on student achievement</u>.

Gage (1972) reports a series of reviews that come to similar conclusions: Committee on the Criteria of Teacher Effectiveness (1953), Wallen and Travers (1963), Getzels and Jackson (1963), and Withall and Lewis (1963).

Conclusions

Our analysis of this literature leads us to three conclusions:

First, the research literature on the relation between teacher

performance and student achievement does not offer an empirical basis for the prescription of teacher-training objectives.

Second, this literature fails to provide such a basis, not because of minor flaws in the statistical analyses, but because of sterile operational definitions of both teaching and achievement, and because of fundamentally weak research designs.

Last, given the well-documented strong association between student achievement and variables such as socio-economic status and race (see Coleman and others, 1966), the effects of techniques of teaching on achievement (as these variables are conventionally defined) are likely to be inherently trivial.

We appreciate the generous support and assistance of Ned A. Flanders and Barak Rosenshine in this investigation.

REFERENCES

American Educational Research Association, Committee on the Criteria of Teache. Effectiveness, 1952. Report of the...Review of Educational Research, 22:238-263.

Anderson, G. J., and H. J. Walberg, 1968. "Classroom Climate and Group Learning." International Journal of Educational Sciences, 2:175-80.

Anthony, B. C. M., 1967. "The Identification and Measurement of Classroom Environmental Process Variables Related to Academic Achievement." Unpublished doctoral dissertation, University of Chicago.

Beiderman, D. D., 1964. "Relationship between Teaching Style and Pupil Behavior." Unpublished doctoral dissertation, University of California at Los Angeles (University Microfilms No. 64-12, 180).

Belgard, M., B. Rosenshine, and N. L. Gage, 1968. "The Teacher's Effectiveness in Explaining: Evidence on Its Generality and Correlation with Pupil's Ratings and Attention Scores," in N. L. Gage et al., Explorations of the Teacher's Effectiveness in Explaining. Stanford: Stanford Center for Research and Development in Teaching, Stanford University (Technical Report No. 4).

Bellack, A. A., H. M. Kleibard, R. T. Hyman, and F. L. Smith, 1966. The Language of the Classroom. New York: Columbia University Press.

Brim, O. G., 1958. Sociology and the Field of Education. New York: Russell Sage Foundation.

Carline, J. L., 1969. "An Investigation of the Relationship between Various Verbal Strategies of Teaching Behavior and Achievement of Elementary School Children." Unpublished doctoral dissertation, Syracuse University.

Chall, J. S., and S. C. Feldmann, 1966. A Study in Depth of First Grade Reading. New York: The City College of the City University of New York (U. S. Office of Education Cooperative Research Project No. 2728).

Coleman, J. S., and others, 1966. Equality of Educational Opportunity. Washington, D. C.: U. S. Government Printing Office.

Conners, C. K., and L. Eisenberg, 1966. The Effect of Teacher Behavior on Verbal Intelligence in Operation Headstart Children. Baltimore: Johns Hopkins University School of Medicine (U. S. Office of Economic Opportunity Headstart Contract No. 510).

Cook, R. E., 1967. "The Effect of Teacher Methodology upon Certain Achievements of Students in Secondary School Biology." Unpublished doctoral dissertation, University of Iowa, Iowa City.

Dubin, R., and T. C. Taveggia, 1968. The Teacher-Learning Paradox: A Comparative Analysis of College Teaching Methods. Eugene, Ore.: Center for the Advanced Study of Educational Administration, University of Oregon.

Elashoff, J. D., 1969. "Analysis of Covariance: A Delicate Instrument." American Educational Research Journal, 6:383-401.

Flanders, N. A., 1970. Analyzing Teaching Behavior. New York: Addison-Wesley Co.

Fortune, J. C., 1966. The Generality of Presenting Behaviors in Teaching Pre-School Children. Memphis, Tenn.: Memphis State University.

-------, 1967. A Study of the Generality of Presenting Behaviors in Teaching. Memphis, Tenn.: Memphis State University (U. S. Office of Education Project No. 6-8468).

-------, N. L. Gage, and R. E. Shutes, 1966. The Generality of the Ability to Explain. Paper presented to the American Educational Research Association, February 1966. Amherst: University of Massachusetts, College of Education.

Furst, N. F., 1967. The Multiple Languages of the Classroom. Paper presented at the meeting of the American Educational Research Association, February 1967. Also available as an unpublished doctoral dissertation, Temple University, Philadelphia, 1967.

Gage, N. L., 1972. Teacher Effectiveness and Teacher Education. Palo Alto, Calif.: Pacific Books.

Getzels, J. W., and P. W. Jackson, 1963. "The Teacher's Personality and Characteristics," in N. L. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally. Pp. 506-82.

Gunnison, J. P., 1968. "An Experiment to Determine the Effects of Changing Teacher Classroom Behavior through Training of Student-Teachers in the Use of the Flanders' Interaction Analysis System." Unpublished doctoral dissertation, Arizona State University, Tempe, Arizona.

Harris, A. J., et al., 1968. A Continuation of the CRAFT Project: Comparing Reading Approaches with Disadvantaged Urban Negro Children in Primary Grades. U. S. Office of Education Project No. 5-0570-2-12-1. New York: Division of Teacher Education of the City University of New York.

Herman, W. L., Jr., et al., 1969. "The Relationship of Teacher-Centered Activities and Pupil-Centered Activities to Pupil Achievement and Interest in 18 First-Grade Social Studies Classes." American Educational Research Journal, 6:227-40.

Hiller, J. H., G. A. Fisher, and W. Kaess, 1969. "A Computer Investigation of Verbal Characteristics of Effective Classroom Learning." American Educational Research Journal, 6:661-77.

Hunter, C. P., 1969. "Classroom Climate and Pupil Characteristics in Special Classes for the Educationally Handicapped." Unpublished doctoral dissertation, University of Southern California, Los Angeles.

Kleinman, G., 1964. "General Science Teacher's Questions, Pupil and Teacher Behaviors, and Pupils' Understanding of Science." Unpublished doctoral dissertation, University of Virginia, Charlottesville, Virginia (University Microfilms No. 65-3961).

LaShier, W. S., Jr., 1967. "The Use of Interaction Analysis in BSCS Laboratory Block Classrooms." Journal of Teacher Education, 18:439-46

Mastin, V. E., 1963. "Teacher Enthusiasm." Journal of Educational Research, 56:385-86.

Medley, D. M., and H. E. Mitzel, 1959. "Some Behavioral Correlates of Teacher Effectiveness." Journal of Educational Psychology, 50:239-46.

Miller, G. L., 1966. "Collaborative Teaching and Pupil Thinking." Journal of Teacher Education, 17:337-58.

Mood, A., 1970. "Do Teachers Make a Difference?" in Do Teachers Make a Difference? A Report on Recent Research on Pupil Achievement. Washington, D. C.: Bureau of Educational Personnel Development, Office of Education. Pp. 1-24.

Morsh, J. E., 1956. Systematic Observation of Instructor Behavior. Development Report No. AFPTRC-TN-56-52. San Antonio: Air Force Personnel and Training Research Center, Lackland Air Force Base.

Nikoloff, S. E. B., 1965. "The Relationship of Teacher Standards to the Written Expression of Fifth and Sixth Grade Children." Unpublished doctoral dissertation, State University of New York at Buffalo, Buffalo, New York (University Microfilms No. 66-3520).

Perkins, H. V., 1964. "A Procedure for Assessing the Classroom Behavior of Students and Teachers." American Educational Research Journal, 1:240-60.

-------, 1965. "Classroom Behavior and Under-Achievement." American Educational Research Journal, 2:1-12.

Pinney, R. H., 1969. "Presentational Behaviors Related to Success in Teaching." Unpublished doctoral dissertation, Stanford University.

Rosenshine, B., 1968. "Objectively Measured Behavioral Predictors of Effectiveness in Explaining." Stanford, Calif.: Technical Report No. 4, Stanford Research and Development Center in Teaching.

-------, and N. F. Furst, 1971. "Research on Teacher Performance Criteria." In B. O. Smith, (Ed.), Research in Teacher Education; A Symposium. Englewood Cliffs, N. J.: Prentice-Hall. Pp. 37-72.

Shutes, R. E., 1969. "Verbal Behaviors and Instructional Effectiveness." Unpublished doctoral dissertation, Stanford University.

M., 1967. The Process of Schooling. New York: Rinehart, and Winston.

Soar, R. S., 1966. An Integrative Approach to Classroom Learning. Philadelphia, Pa.: Temple University (Final Report, Public Health Service Grant No. 5-R11-MH 01096 and National Institute of Mental Health Grant No. 7-R11-MH 02045). (ERIC ED: 033-479).

Solomon, D., W. E. Bezdek, and L. Rosenberg, 1963. Teaching Styles and Learning. Chicago: Center for the Study of Liberal Education for Adults. (ERIC ED: 026-556).

Spaulding, R. L., 1965. Achievement, Creativity, and Self-Concept Correlates of Teacher-Pupil Transactions in Elementary Schools. Hempstead, New York: Hofstra University (U. S. Office of Education Cooperative Research Project No. 1352).

Thompson, G. R., and N. C. Bowers, 1968. Fourth Grade Achievement as Related to Creativity, Intelligence, and Teaching Style. Paper presented to the American Educational Research Association, Chicago, February 1968.

Torrance, E. P., and E. Parent, 1966. Characteristics of Mathematics Teachers that Affect Students' Learning. U. S. Office of Education Cooperative Research Project No. 1020, Contract No. SAE-8993. Minneapolis: Minnesota School Mathematics and Science Center, Institute of Technology, University of Minnesota.

Vorreyer, D. F., 1965. "An Analysis of Teacher Classroom Behavior and Role." Unpublished doctoral dissertation, University of Maryland, College Park.

Walberg, H. J., 1969. "Predicting Class Learning." American Educational Research Journal, 6:529-43.

--------, and G. J. Anderson, 1968. "Classroom Climate and Individual Learning." Journal of Educational Psychology, 59:414-19.

Wallen, N. E., 1966. Relationships between Teacher Characteristics and Student Behavior: Part Three. U. S. Office of Education Cooperative Research Project No. SAE OE 5-10-181. Salt Lake City: University of Utah.

--------, and R. M. W. Travers, 1963. "Analysis and Investigation of Teaching Methods," in N. L. Gage (Ed.), Handbook of Research on Teaching. Chicago: Rand McNally. Pp. 448-505.

Withall, J., and W. W. Lewis, 1963. "Social Interaction in the Classroom," in N. L. Gage (Ed.) Handbook of Research on Teaching. Chicago: Rand McNally. Pp. 683-714.

Wittrock, M., 1962. "Set Applied to Student Teaching." Journal of Educational Psychology, 53:175-80.

Wright, C. J., and G. Nuthall, 1970. "The Relationships between Teacher Behavior and Pupil Achievement in Three Experimental Elementary Science Lessons." American Educational Research Journal, 7:477-91.